Title
A Sing-Song Way of Vocalizing: Generalization and Specificity in Language and Birdsong.

Journal
Neuron, 96(5)

ISSN
0896-6273

Authors
Farias-Virgens, Madza
White, Stephanie A

Publication Date
2017-12-01

DOI
10.1016/j.neuron.2017.11.027

Peer reviewed

molecular process in the downscaling of AMPA receptors, e.g., either as a component to regulate the molecular interaction with AMPA receptors or as a part of signaling pathways. Wang and colleagues lay out a comprehensive and clear picture on how Sema3F mediates the synaptic scaling through modulating the interaction of its holoreceptor complex with and downscaling of AMPA receptors.

Homeostatic plasticity regulated by semaphorin/plexin signaling may also involve the functional coordination of presynaptic and postsynaptic sites. For example, it is well known that retrograde signals from muscle can regulate presynaptic homeostatic plasticity at the neuromuscular synapses in *Drosophila* (Davis and Müller, 2015). Specifically, perturbing muscle activity by attenuating the functions of postsynaptic glutamate receptors leads to presynaptic homeostatic plasticity (i.e., increased neurotransmitter release), which offsets the decreased postsynaptic efficacy and thus restores normal muscle excitation. A recent study published in *Nature* shows that another semaphorin member, Sema2b, which is secreted from muscle, enhances neurotransmitter release from presynaptic motor neurons to maintain normal muscle excitation (Orr et al., 2017). Sema2b is specifically released from muscle and acts on PlexB in motor neurons. The specific action of Sema2b/PlexB is mediated via the potentiation of readily releasable pools of synaptic vesicles through actin dynamics and in turn increases the probability of neurotransmitter release. Thus, during homeostatic plasticity, elevation of activity stimulates the secretion of semaphorins as feedback mechanisms and triggers semaphorin-mediated signaling at synapses to control the pre- and postsynaptic strength through independent molecular mechanisms. The secretion of Sema2b and Sema3F exerts their distinct functions by regulating neurotransmitter release at the presynaptic nerve terminals and postsynaptic neurotransmitter receptors, respectively. The specificity of the semaphorin/plexin signaling in homeostatic plasticity is most likely mediated through the expression and the controlled secretion of specific ligands as well as the coupling of their co-receptors at the pre- and postsynaptic terminals. The characterization of the role of semaphorin/plexin signaling in two different homeostatic plasticity paradigms highlights that the precise control of semaphorin/plexin signaling is critical for maintaining the stability of synaptic transmission and network activity during neural development.

REFERENCES

Davis, G.W., and Müller, M. (2015). Annu. Rev. Physiol. *77*, 251–270.

Kumanogoh, A., and Kikutani, H. (2013). Nat. Rev. Immunol. *13*, 802–814.

Orr, B.O., Fetter, R.D., and Davis, G.W. (2017). Nature *550*, 109–113.

Pozo, K., and Goda, Y. (2010). Neuron *66*, 337–351.

Tran, T.S., Rubio, M.E., Clem, R.L., Johnson, D., Case, L., Tessier-Lavigne, M., Huganir, R.L., Ginty, D.D., and Kolodkin, A.L. (2009). Nature *462*, 1065–1069.

Turrigiano, G.G. (2008). Cell *135*, 422–435.

Turrigiano, G. (2012). Cold Spring Harb. Perspect. Biol. *4*, a005736.

Wang, R., Mellem, J.E., Jensen, M., Brockie, P.J., Walker, C.S., Hoerndli, F.J., Hauth, L., Madsen, D.M., and Maricq, A.V. (2012). Neuron *75*, 838–850.

Wang, Q., Chiu, S.-L., Koropouli, E., Hong, I., Mitchell, S., Easwaran, T.P., Hamilton, N.R., Gustina, A.S., Zhu, Q., Ginty, D.D., Huganir, R.L., and Kolodkin, A.L. (2017). Neuron *96*, this issue, 1084–1098.

A Sing-Song Way of Vocalizing: Generalization and Specificity in Language and Birdsong

Madza Farias-Virgens[1,2] and Stephanie A. White[1,2,*]
[1]Molecular, Cellular, and Integrative Physiology Interdepartmental Program
[2]Department of Integrative Biology and Physiology
University of California, Los Angeles, 610 Charles E. Young Drive East, Los Angeles, CA 90095-7239, USA
*Correspondence: sawhite@ucla.edu
https://doi.org/10.1016/j.neuron.2017.11.027

Spoken languages such as German are extremely discrete, whereas others such as Portuguese are melodic or "sing-song" wherein identifying a word relies on what comes before and after. Perhaps surprisingly, birdsong also exhibits specificity and generalization as articulated by Tian and Brainard (2017).

Efficient performance of motor skills, including ones of vocal nature, relies on the precise control of movements to (1) generate individual motor gestures and (2) rapidly organize them into sequences. Multiple occurrences of the same motor gesture can exhibit substantial variability, adding complexity to their identification by a potential receiver and their articulation by the sender. One source of variability stems from the influence of sequential context, as when gestures present modifications depending on their interaction with other gestures composing a motor sequence. Co-

articulatory effects represent this type of variability, whereby gestures are modified depending on other gestures preceding or following them. Such effects were first described for human speech but have since been expanded to include other complex motor sequences, such as those used to fingerspell American Sign Language (Jerde et al., 2003).

It has therefore been proposed that learning or performing a complex motor skill requires a balance between generalization and specificity. Generalization is defined as the ability to identify or articulate a motor gesture or modification thereof in multiple sequential contexts. Specificity refers to the ability to do so only in a certain sequential context. An example of generalization is the transfer of modifications of a motor gesture induced by external perturbation to contexts other than the learned one. Generalization can thus be efficient for learning, including by reducing the need to relearn how to produce a vocal gesture correctly in each possible context. This adaptive quality of motor gestures is important for the motor response to weakening of muscles, as evidenced in neuromuscular disorders and natural aging.

Human speech generalization can be studied by exposing subjects to perturbations of their own speech patterns, using real-time feedback of transformed speech signals, and analyzing the transfer of adaptations to the same trained phoneme produced in different phonetic contexts, or across phonemes other than the trained one. In a representative early example of this kind of analysis, Houde and Jordan (1998) show that speakers learn to adjust their production of a vowel to compensate for feedback alterations that change the vowel's formants, and that this effect indeed generalizes to the same vowel produced in different words, and across different vowels. A degree of specificity also applies to the identification and articulation of motor gestures, in which modification of phonemes and their associated movements are acquired in a manner that applies only to specific phonetic contexts. This effect has been shown in speech in which subjects receive and efficiently correct for auditory feedback that is altered in opposing ways for the same vowel produced in different words (Rochet-Capellan and Ostry, 2011).

Here, Tian and Brainard present evidence that generalization and specificity of vocal gestures also occur in birdsong, an essential model of learned vocalization that offers the ability to mechanistically test the neural bases of such phenomena. Using Bengalese finches (*Lonchura striata domestica*), the authors first investigate the generalization of learned modifications of vocal gestures (song syllables) by employing a negative reinforcement paradigm in which the bird must learn to shift the pitch of a targeted syllable in a single sequential context to escape a burst of white noise. Previously, investigation of vocal generalization by Hoffmann and Sober (2014) employed a different experimental setting in which Bengalese finches were fitted with miniaturized headphones to selectively perturb auditory feedback of their syllables (Hoffmann and Sober, 2014). Under this training paradigm, the pitch of auditory feedback from a single syllable in each bird's song was perturbed only when it occurred in a specific sequential context. The altered feedback induced the bird to error correct by shifting the pitch of its own syllable in an adaptive direction, i.e., opposite to the auditory feedback. This suggests that songbirds can generate internal error signals specific to vocal gestures, thus paralleling findings in humans who alter the acoustics of single phonemes in response to manipulated auditory feedback (Houde and Jordan, 1998).

In both songbird studies, generalization was tested by measuring the bird's ability to shift the pitch of syllables of the same type as the targeted syllable in sequential contexts other than trained. Findings from both groups demonstrate the same effect, in which learned modifications of the targeted syllable are maximal for the trained sequential context but also partially generalize to other sequential contexts, demonstrating that this phenomenon observed in human behavioral studies applies to birdsong (Houde and Jordan, 1998). They both also attribute a significant influence of sequential factors on the degree of generalization, each one uncovering a different sequence-dependent effect.

In the Hoffmann and Sober (2014) error-correction paradigm, non-targeted syllables changed in the adaptive direction when they were produced in close sequential proximity to the pitch-shifted syllable, suggesting that error information from one syllable is used to generate adaptive vocal changes in other syllables. This finding has been attributed to an "error credit assignment" mechanism by which the sensory error is partially assigned to similar nearby syllables (Wolpert et al., 2011). Interestingly, the same sequence-dependent effect is not reported here by Tian and Brainard. The different outcome may relate to differences in the experimental paradigm and reflect the influence of available sensory cues in generalization. A possible explanation is that the white noise serves as a highly salient negative reinforcement signal the bird uses to localize vocal changes in time, thus assigning errors more specifically to the targeted syllables. Tian and Brainard's white noise paradigm presents us with a newly reported sequence-dependent effect. They show that the magnitude of generalization for a given non-targeted context can be explained by its sequential similarity to the targeted context, which they call "contextual similarity."

Using the same white noise negative reinforcement paradigm as they did for generalization, Tian and Brainard go further to directly interrogate the specificity of the learned modifications: they measure the bird's ability to shift the pitch of a targeted syllable in one direction for a single sequential context, and in the opposite direction for another sequential context. Again, their findings mirror human behavioral studies in which opposite modifications of a targeted gesture can be learned for different sequential contexts (Rochet-Capellan and Ostry, 2011). Tian and Brainard's paradigm thus directly compares, in the same study, generalization and specificity in songbird vocal learning. Their findings endorse birdsong as a tractable model for investigating components of human speech and firmly establish the Bengalese finch in the study of context-dependent vocal motor learning.

The neural mechanisms underlying birdsong have been studied for >100 years, providing us with a well-characterized picture of a system composed of anatomically distinct groups of neurons within an anterior forebrain pathway (AFP) and a posterior motor pathway

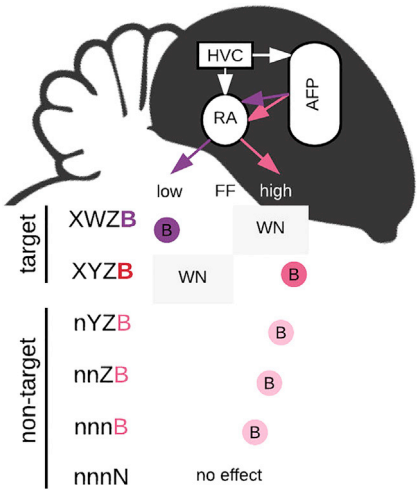

Figure 1. Biasing Signals from Songbird Cortico-Basal Ganglia Circuitry

From top to bottom: biasing signals from the avian cortico-basal ganglia circuitry (AFP) are sent to motor nucleus RA. These can generate changes in the pitch of a syllable (e.g., "B") targeted by white noise (WN) in one direction (purple arrow, low) for a certain sequential context or in the opposite direction for another sequential context (pink arrow, high). The magnitude of generalization of pitch changes to syllables of the same type as the targeted syllable (light pink) is relative to how many syllables are shared between target and non-target contexts. ("X," "Y," and "Z"; "n" denotes any other syllable type). FF, fundamental frequency.

(Bolhuis et al., 2010). The AFP consists of a cortico-basal ganglia loop and intersects with the posterior motor pathway, which itself consists of direct projections from cortical neurons onto brainstem motor neurons that control the vocal organs. Tian and Brainard take advantage of this prolific body of work to test key hypotheses regarding the neural mechanisms that operate to support the coding or representation of vocal gestures as it relates to generalization and specification. Specifically, it has been proposed that modifications of targeted syllables are generated by biasing signals from the AFP. This hypothesis is based on a collection of more recent studies, using a variety of experimental methods, including chemical inactivation, electrical stimulation, and single-unit neural recordings, that demonstrate the role that the AFP plays in natural vocal variability and in responding to perceived vocal errors, which can be used for learning in juvenile and adult birds (Woolley and Kao, 2015). The authors further hypothesize that if those same mechanisms contribute to all adaptive modifications of song, they could expect that the early expression of learning in the target context would also rely on biasing signals from the AFP.

Their results indeed show that chemical inactivation of LMAN (lateral magnocellular nucleus of the nidopallium), the cortical output nucleus of the AFP, reduces the magnitude of expressed learning that is specific to the trained sequential context, an effect that is reversible upon wash-out of the drug. Tian and Brainard propose a model to integrate their observations. They suggest that the AFP provides context-specific biasing signals that modify a context-independent syllable representation in the motor circuitry involving RA (robust nucleus of the arcopallium), another cortical nucleus in the songbird brain that is analogous to the human primary motor cortex (Pfenning et al., 2014; Figure 1). They provide evidence supporting their model's prediction that generalization would be reduced when incongruent modifications of a syllable are reinforced in different contexts. This would not be the case if separate representations of a given syllable in each context were being generated and consolidated.

The overall architecture of speech-related areas in human brain finds parallels in the organization of the AFP in the songbird brain (Bolhuis et al., 2010). In both humans and songbirds, basal ganglia circuitry is highly interconnected with widespread areas of the cortex, including areas related to several aspects of vocal behavior, such as vocal motor control, auditory processing, and sensory-motor integration. This suggests that the cortico-basal ganglia circuitry not only influences motor aspects of vocal processes, but also plays an important role in other related cognitive functions, including auditory categorization. In speech category learning, this idea is supported by neuroimaging evidence that demonstrates the recruitment of the basal ganglia learning system during auditory categorization tasks that are based on outcome feedback (Lim et al., 2014). Learning appropriate behavioral actions to achieve goals in specific environments can be thought of as a reinforcement process in which one builds and updates predictions about receiving future rewards. Based on such predictions, behavior adjusts adaptively to maximize future rewards such that actions leading to rewards are reinforced, whereas incorrect behaviors leading to punishment or no rewards are modified. This overarching model provides a useful framework for understanding the way by which cortico-basal ganglia circuitry contributes to the generation of context-specific biasing signals necessary also for auditory/vocal learning. While the exact nature of auditory representation remains for the most part unexplained, in general, evidence from both birdsong and human speech agree in which they implicate cortico-basal ganglia circuitry in eliciting biasing signals that are likely related to auditory categorization in goal-oriented contexts, and are necessary to guide responses to perceived vocal errors in adaptive ways.

REFERENCES

Bolhuis, J.J., Okanoya, K., and Scharff, C. (2010). Nat. Rev. Neurosci. *11*, 747–759.

Hoffmann, L.A., and Sober, S.J. (2014). J. Neurosci. *34*, 5564–5574.

Houde, J.F., and Jordan, M.I. (1998). Science *279*, 1213–1216.

Jerde, T.E., Soechting, J.F., and Flanders, M. (2003). J. Neurosci. *23*, 2383–2393.

Lim, S.J., Fiez, J.A., and Holt, L.L. (2014). Front. Neurosci. *8*, 230.

Pfenning, A.R., Hara, E., Whitney, O., Rivas, M.V., Wang, R., Roulhac, P.L., Howard, J.T., Wirthlin, M., Lovell, P.V., Ganapathy, G., et al. (2014). Science *346*, 1256846.

Rochet-Capellan, A., and Ostry, D.J. (2011). J. Neurosci. *31*, 2657–2662.

Tian, L.Y., and Brainard, M.S. (2017). Neuron *96*, this issue, 1168–1177.

Wolpert, D.M., Diedrichsen, J., and Flanagan, J.R. (2011). Nat. Rev. Neurosci. *12*, 739–751.

Woolley, S.C., and Kao, M.H. (2015). Neuroscience *296*, 39–47.